

10/A
CIA
11/24/03

Amendments to the Claims:

This listing of claims replaces all prior versions and listings of claims in the application:

Listing of Claims:

1. (Currently Amended) A method of expanding an effective active vocabulary of a speech recognition system, the method comprising:

using a speech recognizer to perform speech recognition on a user utterance to produce one or more recognition candidates, the speech recognition comprising comparing digital values representative of the user utterance to a set of acoustic models representative of an active vocabulary of the system, the set of acoustic models including models of words and models of word fragments,

receiving the recognition candidates from the speech recognizer, and
when a received recognition candidate includes a word fragment:

determining whether the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary of the speech recognition system, wherein forming the proposed word includes using a spelling rule associated with the word fragment that causes the spelling of the proposed word to differ from a spelling that would result from merely concatenating the particular word fragment with the one or more adjacent word fragments or words;

if the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary of the speech recognition system, modifying the recognition candidate to substitute the proposed word for the word fragment and the one or more adjacent word fragments or words used to form the proposed word; and

if the word fragment may not be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary of the speech recognition system, discarding the recognition candidate.

2. (Original) The method of claim 1, wherein the expanded effective vocabulary comprises words from the backup dictionary that are formed from a combination of words and word fragments or word fragments and word fragments from an active vocabulary that includes words and word fragments, and words from the active vocabulary.

3. (Original) The method of claim 1, wherein word fragments comprise suffixes, prefixes, and roots that are not words.

AI 4. (Original) The method of claim 3, wherein:
one or more spelling rules are associated with each prefix and each suffix,
determining whether the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word comprises using a prefix or suffix as the particular word fragment and using an associated spelling rule in forming the proposed word, and

as a result of using the associated spelling rule, a spelling of the proposed word differs from a spelling that would result from merely concatenating the particular word fragment with the one or more adjacent word fragments or words.

5. (Original) The method of claim 4, wherein determining whether the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word comprises:

retrieving from the received recognition candidate a sequence that includes the particular word fragment and adjacent word fragments or words; and
determining if the sequence is a valid sequence.

6. (Original) The method of claim 5, wherein a valid sequence includes only one or more allowed adjacent combinations of word fragments and words.

7. (Original) The method of claim 6, wherein allowed adjacent combinations comprise one or more prefixes, followed by a root or a word, followed by one or more suffixes; a root or a word followed by one or more suffixes; and one or more prefixes followed by a root or a word.

8. (Original) The method of claim 6, wherein allowed adjacent combinations comprise one or more prefixes, followed by one or more roots or words, followed by one or more suffixes; one or more roots or words followed by one or more suffixes; and one or more prefixes followed by one or more roots or words.

AI 9. (Original) The method of claim 4, further comprising combining the particular word fragment with the one or more adjacent word fragments or words to form a second proposed word that differs from the first proposed word by using a second associated spelling rule in forming the proposed word.

10. (Cancelled)

11. (Original) The method of claim 1, wherein determining whether the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary of the speech recognition system comprises searching the backup dictionary for the proposed word.

12. (Original) The method of claim 1, wherein modifying the recognition candidate comprises:

forming a prospective recognition candidate by modifying the recognition candidate to substitute the proposed word for the word fragment and the one or more adjacent word fragments or words used to form the proposed word; and

if the prospective recognition candidate includes an additional word fragment:

further processing the prospective recognition candidate to generate an additional word using the additional word fragment and one or more adjacent words or word fragments, and forming a final recognition candidate by replacing the additional word fragment and the one or more adjacent words with the additional word.

13. (Original) The method of claim 1, wherein a score is associated with the received recognition candidate, the method further comprising producing a score associated with the modified recognition candidate by rescoring the modified recognition candidate.

Al 14. (Original) The method of claim 13, wherein:
the score associated with the received recognition candidate includes an acoustic component and a language model component, and
rescoring the modified recognition candidate comprises generating a language model score for the modified recognition candidate.

15. (Original) The method of claim 14, wherein producing the score associated with the modified recognition candidate comprises combining the acoustic component of the score for the received recognition candidate with the language model score generated for the modified recognition candidate.

16. (Original) The method of claim 14, wherein rescoring the modified recognition candidate comprises generating an acoustic model score for the modified recognition candidate and producing the score associated with the modified recognition candidate comprises combining the acoustic model score generated for the modified recognition candidate with the language model score generated for the modified recognition candidate.

17. (Original) The method of claim 13, wherein:
the score associated with the received recognition candidate includes an acoustic component and a language model component; and
rescoring the modified recognition candidate comprises generating an acoustic score for the modified recognition candidate.

Al 18. (Original) The method of claim 17, wherein producing the score associated with the modified recognition candidate comprises combining the language model component of the score for the received recognition candidate with the acoustic score generated for the modified recognition candidate.

19. (Original) The method of claim 1, further comprising generating the acoustic model of a word fragment by:

comparing a word of the active vocabulary to a similar word of a backup dictionary to identify a word fragment that may be used to convert the word of the active vocabulary to the word of the backup dictionary, and

generating the acoustic model of the word fragment using a portion of an acoustic model of the word of the backup dictionary that is not included in an acoustic model of the word of the active vocabulary.

20. (Original) The method of claim 19, wherein comparing a word of the active vocabulary to a similar word of a backup dictionary comprises comparing spellings of the two words.

21. (Original) The method of claim 1, further comprising generating the acoustic models of word fragments by:

comparing words of the active vocabulary to similar words of a backup dictionary to identify spelling rules that may be used to convert the words of the active vocabulary to words of the backup dictionary; and

employing the spelling rules in identifying word fragments.

22. (Original) The method of claim 21, wherein employing the spelling rules in identifying the word fragments comprises:

grouping spelling rules together to form possible affixes, the affixes including prefixes and suffixes; and

analyzing words of the backup dictionary using the affixes to identify roots that may be combined with the affixes to produce words of the backup dictionary.

23. (Original) The method of claim 22, further comprising generating acoustic models for the roots using portions of acoustic models of the words of the backup dictionary that are not included in acoustic models of the affixes.

24. (Original) The method of claim 22, further comprising adding affixes and roots to the active vocabulary as word fragments.

25. (Original) The method of claim 24, further comprising storing a set of spelling rules in association with an affix in the active vocabulary.

26. (Original) The method of claim 24, further comprising creating a language model associated with the active vocabulary.

27. (Original) The method of claim 26, wherein creating the language model comprises:

retrieving a training collection of text, the training collection of text comprising words from the backup dictionary and words from the active vocabulary;

modifying the training collection of text by replacing any splittable backup dictionary words with their corresponding words and word fragments; and

generating language model scores for words and word fragments of the active vocabulary using the modified collection of text.

28. (Original) The method of claim 26, wherein creating the language model comprises creating an N-gram language model.

AI 29. (Original) The method of claim 28, wherein creating the N-gram language model comprises:

retrieving a training collection of text, the training collection of text comprising words from the backup dictionary and words from the active vocabulary;

determining a frequency of each N-gram word sequence that appears in the training collection of text;

modifying the N-gram word sequences by replacing any splittable backup dictionary words with their corresponding words and word fragments;

based on the N-gram word sequence frequencies, determining a frequency of each modified N-gram sequence that includes words, word fragments, or words and word fragments; and

based on the N-gram word and word fragment sequence frequencies, generating an N-gram language model for the words and word fragments of the active vocabulary.

30. (Original) The method of claim 1, wherein word fragments comprise syllables, and each syllable includes a unit of spoken language.

31. (Original) The method of claim 30, wherein the unit of spoken language comprises a single uninterrupted sound formed by a vowel or diphthong alone, of a syllabic consonant alone, or of either with one or more consonants.

32. (Original) The method of claim 30, wherein a syllable comprises a vowel and one or more consonants clustered around the vowel.

33. (Original) The method of claim 1, wherein determining whether the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in the backup dictionary of the speech recognition system comprises searching the backup dictionary for the proposed word based on a pronunciation of the proposed word.

34. (Currently amended) A method of recognizing speech, the method comprising:
using a speech recognizer to perform speech recognition on a user utterance to produce a set of one or more recognition candidates, the speech recognition comprising comparing digital values representative of the user utterance to a set of acoustic models representative of an active vocabulary of the system, the set of acoustic models including models of words, models of roots that are not words, and models of affixes that are not words, the affixes including prefixes and suffixes,

receiving the recognition candidates from the speech recognizer, and
when a received recognition candidate includes an affix:

combining the affix with one or more adjacent words, roots, or other affixes to form a new word, wherein forming the new word includes using a spelling rule associated with the affix that causes the spelling of the new word to differ from a spelling that would result from merely concatenating the affix with the one or more adjacent words, roots, or other affixes; and

modifying the recognition candidate to substitute the new word for the affix and the one or more adjacent words, roots, or other affixes used to form the new word.

35. (Original) A method of generating an acoustic model of a word fragment, the method comprising:

comparing a word of an active vocabulary to a similar word of a backup dictionary to identify a word fragment that may be used to convert the word of the active vocabulary to the word of the backup dictionary, and

generating the acoustic model of the word fragment using a portion of an acoustic model of the word of the backup dictionary that is not included in an acoustic model of the word of the active vocabulary.

36. (Original) The method of claim 35, wherein comparing a word of the active vocabulary to a similar word of a backup dictionary comprises comparing spellings of the two words.

37. (Original) A method of generating acoustic models of word fragments, the method comprising:

comparing words of an active vocabulary to similar words of a backup dictionary to identify spelling rules that may be used to convert the words of the active vocabulary to words of the backup dictionary; and

employing the spelling rules in identifying word fragments.

38. (Original) The method of claim 37, wherein employing the spelling rules in identifying word fragments comprises:

grouping spelling rules together to form possible affixes, the affixes including prefixes and suffixes; and

analyzing words of the backup dictionary using the affixes to identify roots that may be combined with the affixes to produce words of the backup dictionary.

39. (Original) The method of claim 38, further comprising generating acoustic models for the roots using portions of acoustic models of the words of the backup dictionary that are not included in acoustic models of the affixes.

40. (Original) The method of claim 38, further comprising adding affixes and roots to the active vocabulary as word fragments.

Al 41. (Original) The method of claim 40, further comprising storing a set of spelling rules in association with an affix in the active vocabulary.

42. (Original) The method of claim 40, further comprising creating a language model associated with the active vocabulary.

43. (Original) The method of claim 42, wherein creating the language model comprises:

retrieving a training collection of text, the training collection of text comprising words from the backup dictionary and words from the active vocabulary;

modifying the training collection of text by replacing any splittable backup dictionary words with their corresponding words and word fragments; and

generating language model scores for words and word fragments of the active vocabulary using the modified collection of text.

44. (Original) The method of claim 42, wherein creating the language model comprises creating an N-gram language model.

45. (Original) The method of claim 44, wherein creating the N-gram language model comprises:

retrieving a training collection of text, the training collection of text comprising words from the backup dictionary and words from the active vocabulary;

determining a frequency of each N-gram word sequence that appears in the training collection of text;

modifying the N-gram word sequences by replacing any splittable backup dictionary words with their corresponding words and word fragments;

based on the N-gram word sequence frequencies, determining a frequency of each modified N-gram sequence that includes words, word fragments, or words and word fragments; and

based on the N-gram word and word fragment sequence frequencies, generating an N-gram language model for the words and word fragments of the active vocabulary.

46. (Currently amended) A computer-implemented speech recognition system that uses an expanded effective active vocabulary, the system comprising:

a storage device configured to store an active vocabulary that includes multiple entries corresponding to words, commands, and word fragments; and

a processor configured to:

receive data representing a user utterance,

produce one or more recognition candidates, by comparing digital values representative of the user utterance to a set of acoustic models representative of the active vocabulary of the system, the set of acoustic models including models of words and models of word fragments,

when a produced recognition candidate includes a word fragment:

determine whether the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary of the speech recognition system, wherein forming the proposed word includes using a spelling rule

associated with the word fragment that causes the spelling of the proposed word to differ from a spelling that would result from merely concatenating the particular word fragment with the one or more adjacent word fragments or words;

if the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary, modify the recognition candidate to substitute the proposed word for the word fragment and the one or more adjacent word fragments or words used to form the proposed word; and

if the word fragment may not be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary of the speech recognition system, discard the recognition candidate.

47. (Original) The system of claim 46, wherein the expanded effective active vocabulary comprises words from the backup dictionary that are formed from a combination of words and word fragments or word fragments and word fragments from an active vocabulary that includes words and word fragments, and words from the active vocabulary.

48. (Original) The system of claim 46, wherein the processor determines whether the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in the backup dictionary by searching the backup dictionary for the proposed word.

49. (Original) The system of claim 46, wherein the processor modifies the recognition candidate by:

forming a prospective recognition candidate by modifying the recognition candidate to substitute the proposed word for the word fragment and the one or more adjacent word fragments or words used to form the proposed word; and

if the prospective recognition candidate includes an additional word fragment:

further processing the prospective recognition candidate to generate an additional word using the additional word fragment and one or more adjacent words or word fragments, and forming a final recognition candidate by replacing the additional word fragment and the one or more adjacent words with the additional word.

50. (Original) The system of claim 46, wherein a score is associated with the received recognition candidate, and the processor is further configured to produce a score associated with the modified recognition candidate by rescoring the modified recognition candidate.

Al 51. (Currently amended) Computer software, residing on a computer readable medium, for a speech recognition system that uses an expanded effective active vocabulary to recognize words, and commands, the computer software comprising instructions for causing a computer to perform the following operations:

receive data representing a user utterance,
produce one or more recognition candidates, by comparing digital values representative of the user utterance to a set of acoustic models representative of an active vocabulary of the system, the set of acoustic models including models of words and models of word fragments,
when a produced recognition candidate includes a word fragment:

determine whether the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary of the speech recognition system, wherein forming the proposed word includes using a spelling rule associated with the word fragment that causes the spelling of the proposed word to differ from a spelling that would result from merely concatenating the particular word fragment with the one or more adjacent word fragments or words;

if the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary, modify the recognition

candidate to substitute the proposed word for the word fragment and the one or more adjacent word fragments or words used to form the proposed word; and

if the word fragment may not be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary of the speech recognition system, discard the recognition candidate.

52. (Original) The computer software of claim 51, wherein the expanded effective active vocabulary comprises words from the backup dictionary that are formed from a combination of words and word fragments or word fragments and word fragments from an active vocabulary that includes words and word fragments, and words from the active vocabulary.

53. (Original) The computer software of claim 51, wherein determining whether the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in the backup dictionary comprises searching the backup dictionary for the proposed word.

54. (Original) The computer software of claim 51, wherein modifying the recognition candidate comprises:

forming a prospective recognition candidate by modifying the recognition candidate to substitute the proposed word for the word fragment and the one or more adjacent word fragments or words used to form the proposed word; and

if the prospective recognition candidate includes an additional word fragment:

further processing the prospective recognition candidate to generate an additional word using the additional word fragment and one or more adjacent words or word fragments, and

forming a final recognition candidate by replacing the additional word fragment and the one or more adjacent words with the additional word.

Applicant : Jonathan H. Young et al.
Serial No. : 09/390,370
Filed : September 7, 1999
Page : 16 of 27

Docket No.: 13865-023001

55. (Original) The computer software of claim 51, wherein a score is associated with the received recognition candidate, the computer software comprising instructions for causing the computer to produce a score associated with the modified recognition candidate by rescoreing the modified recognition candidate.
